Eliminativism without Tears[1]

Abstract: The paper begins by showing why, given the findings of neuroscience, brain states don't have propositional content. It then examines a leading attempt to attribute content to brain states owing to their functional, i.e. evolutionary role, and show why it is best viewed as a *modus tollens* argument against brain states having content. The paper goes on to try to dispose of a prominent argument that consciousness is the source of proposition content, owing to its role in conferring intentionality on thought. Finally, it attempts to draw the force of a powerful argument that suggests eliminativism's denial that thought has content is self-refuting.

Introduction

The progress of neuroscience will eventually force philosophy to adopt eliminativism. It has already advanced enough to force the philosophy of mind to take eliminativism seriously. But eliminativism is widely held to be incoherent. Therefore, the advance of neuroscience makes eliminativism's apparent incoherence every philosopher's problem. In this paper I show why neuroscience bids fair to vindicate eliminativism, and try to show how its incoherence problem can be mitigated.

Eliminativism is the thesis that the brain does not store information in the form of unique[2] sentences that express statements or propositions[3] or anything like them.[4] It

---

[1] Thanks to Walter Sinnot–Armstrong, Fred Dretske, Gordon Steenbergen, Daniel Kraemer, Owen Flanagan and Bavid Barack for comments, and suggestions. Naturally, they don't agree with almost any of this.

[2] The qualification 'unique' is not very strong, but probably useful to avoid some irrelevant objections. If a neural circuit represents a finite disjunction of sentences, statements or propostions, then it represents their unique disjunction. If the disjunction is not finite, then it cannot contain the disjunction of them. Cf footnote 2. If there is a always a set of multple sentences, statements or propositons any one of which is an equally good candidate for being stored in a neural circuit, and it is logically impossible to identify any one of them as better than the others, then it's safe to say the neural circuit contains at most their disjunction.

[3] Of course individuating sentences is a very different matter from individuating propositions. Neural circuits may contain propositions without containing sentences in some language of thought, but there will have to be some aspects of neural circuitry which token—i.e. physically express—the propositions they contain. For convenience I'll refer to these tokens as sentences, without committing myself to a full blown langauge of thought. Since some sort of tokens will be required to express propositions, unless the tokens do so recursively, a neural cirucit can't encode an infinite number of distinct propositions.

denies the intentionality of thought.[5] Eliminativism does not deny the existence of consciousness or qualitative aspects of experience. It does deny that they are sources or evidence for the intentionality of thought.

The problem of incoherence that this thesis faces is easy to state if we employ a distinction made familiar by John Searle [1980, 1983], that of 'derived' v 'original' intentionality. Public language has 'derived' intentionality. The noises we make and the marks we produce have intentional content owing to their causes. They are symbols, have meaning, express sentences, statements, propositions, because they result from processes that involve brain states with 'original' intentionality. By causing our speech and writing, the brain states that have original intentional content confer derived intentionality on them.

Eliminativism holds that there is no original intentionality. Without it there is no derived intentionality and so our speech and writing have no meaning, they are merely noises and chicken tracks. Without original intentionality no one can think *about* anything and no noise or mark they make can have derived intentionality; no noise or mark can be about anything or a symbol of anything. *Ergo*, the thesis of eliminativism cannot be expressed in speech or writing, and it cannot be thought either. If eliminativism is true then we cannot have the thought that it is true or express that thought in speech or writing. Eliminativism is incoherent.[6]

The defense of eliminativism advanced here adopts Searle's distinction and accepts the premise that speech and writing have at most only derived intentionality; the only way they can have it is owing to the original intentionality, if any, of the brain states of speakers and writers, listeners and readers, who interpret speech and writing as symbols, not merely as signs.

The defense to be mounted makes two further perhaps important but uncontroversial assumptions: First, the brain acquires, stores, uses and transmits information. In fact it stores a vast quantity of it. Eliminativism accepts that there is a great deal of information transmitted, stored and employed in nature. This is the sort of information that has been under discussion in philosophy at least since Dretske (1981). Second, it assumes that much of the information the human brain acquires and transmits comes to it and leaves it via speech—noises coming out of people's mouths, signs made by their bodies (usually their hands) and writing--marks, inscriptions, print, and more

---

[4] Eliminativism's denial of propositional content extends to the denial neural circuitry contains information in the form of distinct names and verbs, subjects and predicates, topics and comments, to use Dretske's terminology (1988) or anything else that would make them truth-apt.

[5] Eliminativism is thus a stronger thesis than the denial that the kinds of folk psychology are natural ones. It denies that any intentional kinds are natural.

[6] This is a far more serious problem for eliminativism than the alleged pragmatic contradiction of 'believing that there are no propositional attitudes,' since beliefs are, by definition, propositional attitudes. There are a variety of alternative dispositional accounts of belief available. The real problem for eliminativism is its denial that there is anything in the brain or elsewhere that qualifies as carrying truth values.

recently pixels. Eliminativism is the thesis that all this happens without the noises or marks having derived intentionality, and without the brain states having original intentionality. For example eliminativism doesn't deny that you are acquiring information (or perhaps misinformation) by the words you are reading. It only denies that the words you are now reading are symbols, have meaning, express propositions, have truth-values.

The next section of this paper shows why neuroscience makes eliminativism about propositional attitudes unavoidable. The following section explores the powerful argument for eliminativism that stems from teleosemantics. This leaves conscious experience as a possible source for intentionality. But, as a source for intentional content, conscious experience has already been excluded by neuroscience. The final part develops as much of a solution as required to avoid the charge of incoherence that faces eliminativism.

1. Why brain states don't have propositional content

Many areas of recent advance in neuroscience are converging on the conclusion that neural circuitry does not record, store or transmit information in forms that could express propositions. The most convincing for our purposes is its understanding of memory, in particular what neuroscience calls "implicit" or "nondeclarative" memory—skills, abilities, conditioning, and 'declarative" or "explicit" memory, which it further divides into "episodic" memory—information about past events personally experienced and "semantic memory,"—information about general facts. Declarative or explicit memories are the ones we ordinarily suppose to be propositional. The implicit/explicit or nondeclarative/declarative distinction mirrors an epistemic distinction introduced by Ryle (1949) between knowledge how and knowledge that. The seats of these two types of memories appear to be separated in the brain. Explicit memory is subserved by structures in the temporal lobe of the cerebrum (and especially the hippocampus and neocortex), while implicit learning involves learning processes in the sensory-motor pathways of organisms, including invertebrates that do not have anything like a cerebrum.

The neural anatomy of the brain is comprised of 10^{11} neurons, each synapsing with up to a 1000 other neurons. Almost all of these neurons do only one thing, moving a relatively small number of molecules, mainly potassium, sodium, and chlorine to other neurons. They do this largely by using a small number of neurotransmitters that open and close channels through which the molecules move between the neurons. The only neurons that don't exactly work this way are the ones that respond directly to sensory inputs, and the ones that affect muscle fibers. But even these neurons connect to other neurons in the same way. The potassium, sodium and chlorine molecules are charged, and their movement conveys electrical potentials. When a sufficiently large number of electrical signals are sent through a neuron over a short enough period, a causal chain from the potassium, sodium and chlorine molecules to the DNA in the nuclei of the neuron switches on certain genes, whose protein products build new synaptic connections that strengthen signal transmission. That's all there is to the brain: a vast number of input/out put circuits (and input/output circuits composed of input output circuits) all pretty much the same in their molecular neurobiology. And it is the in/put circuits that carry all the information the brain does. Such circuitry cannot carry this information around semantically. That is, it doesn't carry it around in sentence-tokens or statements

that have truth-values or component parts that are anything like nouns and verbs with referents and meanings.

How can neuroscience be confident that this is the case? The specific findings that force this conclusion on neuroscience were made over a 30 or 40-year period largely by Eric Kandel in research that eventuated in a Nobel Prize. Much of this work is reported in Bailey, Bartsch, and Kandel (1996). Kandel and his coworkers began by figuring out how classical conditioning produces changes in the neurons that store a learned response, a capacity, disposition or ability to respond to a stimulus. This work, revealing the neural basis of implicit--skill/ability—memory, employed the sea slug, *Aplasia Californica*, whose neural circuitry is accessible and simple. Implicit memory in the sea slug come in distinct short term and long term versions, and both of these are dependent on the number of training trials to which the neural circuits are exposed. Owing to advances in neurogenomics—the use of knockout and gene silencing techniques in the study of neurons--the macromolecular differences between short and long term implicit memory discovered in *Aplasia*, were extended to understanding implicit memory in *C.elegans*, and *Drosophila*.

In all three species, these studies reveal unsurprisingly enough that the difference between short-term and long term memory is a fairly obvious anatomical difference: short term memory is a matter of establishing temporary bonding relationships between molecules in the synapses that degrade quickly, while long term implicit memory results from building new synaptic structures. The former is called 'short term potentiation,' the latter is 'long term potentiation' or LTP.

Short-term implicit learning results from conditioning in which a chain of molecular signals and ambient catalytic molecules produce a short-lived modification in the concentration and the confirmation (secondary and tertiary structure or shape which changes binding and/or catalytic activity) of neurotransmitter-molecules in preexisting synapses. The neural pathway has 'remembered' how to respond to the stimulus. Long-term implicit memory appears to be mainly the result of the stimulation of somatic genes to orchestrate the production of ***new*** synapses connecting sensory and motor neurons. In long-term implicit memory the initial steps are the same as in short-term learning. But something else happens: some of the larger number of molecules (that result from repeated stimulation) diffuse to the sensory neuron's nucleus where they switch genes whose molecular products form new synaptic connections between the sensory neurons and the motor neurons. Long-term implicit memory is realized by an anatomical change at the cellular level that involves switching on a gene that produces more synaptic connections. Each of the new synaptic connections work in the same way as the smaller number of connections laid down for short term implicit memory, but their larger number means that the learned response will be manifested even if a significant number of the synaptic connections degrade, as happens over time. Thus, the new construction of additional synaptic connections provides for long-term implicit memory.

What about explicit, declarative memory, composed of "semantic memory" and episodic memory—what can be expressed in propositions?

Explicit memories storage is localized to the temporal lobe, initially in the hippocampus, and then (shifted by "consolidation") in the neocortex, structures unknown in the sea slug, the worm for the fruit fly.

Studies of neural processing in these regions of the temporal lobe began with the determination that neural pathway there are subject to long term potentiation, LTP, the process in which synapses become much more sensitive when they are repeatedly stimulated. In particular repeated stimulation by the same concentrations of neurotransmitters of the neurons in the hippocampus results in much higher production by them of neurotransmitters that stimulate down-stream neurons.

Kandel et al.'s [1996] studies of LTP showed that the ***same*** molecular mechanisms, involving the same somatic genes that build new synaptic connections in the Aplasia implicit long-term memory, are responsible for all the forms of LTP in all the hippocampal pathways that subserve explicit memory in vertebrates.[7] The same genes build the same new anatomical structures in both long term explicit and long term implicit memory.

It's not just information storage in the neural circuitry of the hippocampus that is the same in its structure as implicit memory in the sea slug. Declarative or explicit memories—propositional knowledge—are in fact moved from the hippocampus to information storage circuitry in the neocortex (a process known as "consolidation"). Both the process of distributing various kinds of information from the hippocampus to visual, auditory, parietal cortices, and the storage in these parts of the brain are the result of and consist in the same molecular and neurogenomic modifications of neural circuitry as Kandel discovered in the sea slug when it acquires long term implicit memories—abilities, capacities, dispositions to respond to stimuli.

The molecular biology of long term implicit memory in the sea slug and long term explicit memory in the human appears also to be substantially the same, indeed identical except for some molecular differences that don't effect the configuration of the neurotransmitters and the nucleic acid sequence difference of the genes and RNAs that regulate changes in the micro-architecture of synaptic connections. The details of the

---

[7] They write,

> Similar to the presynaptic facilitation in *Aplasia*, both mossy fiber and Schaffer collateral LTP [two of the three types of LTP in mammalian hyppocampi and neocortex] have distinct temporal phases…The early phase is produced by a single titanic stimulation [release of neurotransmitters], lasts 1-3 hours, and requires only covalent modification of preexisting proteins. By contrast, the late phase is induced by repeated titanic stimulations, and is dependent on new proteins and RNA synthesis. As is the case with long-term memory in *Aplasia*, on the cellular level there is a consolidation switch, and the requirement for [gene] transcription in LTP has a critical time window. In addition, the late transcription-dependent phase of TLP is blocked by inhibitors of **PKA** … Recent studies by Nguyen and Kandel now indicate that these features of LTP also apply to a third major hippocampal pathway, the medial performant pathway…Thus, as in *Aplasia* presynaptic facilitation, cAMP-mediated transcription appears to be the common mechanism for the late form of LTP in all three pathways within the hippocampus. [p. 13452]

neural connections that constitute long term storage of implicit memory—storage of dispositions and abilities, what Ryle called knowledge how—differ only by number of connections from long-term storage of explicit memory—"declarative memory," what we think of as propositional knowledge. But if long-term explicit memory storage differs only in degree, just bigger and bigger input/output circuits) from long-term implicit memory storage, what looks like propositional knowledge is nothing but a large number of synaptic connections each one of which is a bit of associative learning, a neural circuit that realizes a conditional disposition to respond to stimuli in an environmentally appropriate way, a little bit of knowledge how.

Recall the point that the sensory-motor pathway produced by classical conditioning in the *Aplasia*, constitutes the stored disposition to respond to noxious or positively rewarding stimuli in an environmentally appropriate manner. Move the same circuits to the hippocampus, multiply their numbers by several orders of magnitude, and the result is long term explicit memory, what Kandel called "declarative" because in humans the information stored can often be recalled "at will," and when it is recalled, it can be verbalized. Now, assume what neuroscientists and all other life scientists must: *Natura non facit saltum.* Differences in the number, location, and wiring of individual neural circuits can only turn them from small sets of input/output systems into larger ones. It can't turn them from one kind of thing—the stimulus/response wiring of a sea slug into an entirely different kind of thing, stored sentential content in the neurons.[8] This seems to be a conclusion vouched safe by two common sense observations: the ability to ride a cycle cannot be adequately captured by any number of propositions about bike riding or about anything else for that matter; no proposition about how the world is arranged can be identical to any set of dispositions or abilities on the part of someone who believes it. Of course eliminativism is not going to rely very strongly on such observations.

If any of the research programs of cognitive neuroscience pan out, we will discover higher levels of organization in the brain, programs that operate on populations of thousands or millions of neurons that store information as input/output circuitry. There may even be systems or programs that operate on formal properties of these sets of neural circuits to manifest some of the infinitary and recursive capacities thought reflects in behavior—e.g., speech, or mathematical calculation. But whatever neural structures these higher levels of organization operate on, they won't be ones that store information in sentences, they won't be ones that can be combined together by any concatenation or wiring to constitute larger structures that do encode information sententially. And as Kandel won a Nobel Prize for showing, they don't need to do so for the brain to store the information it has.[9]

---

[8] Of course, larger neural circuits can have functions that their component circuits dont have, and these could accord such larger circuits content. The next section takes up this teleomantic strategy. There is no scope for greater organization or complexity among neural circuits to give rise to some emergent content. In the brain organization and complexity is just a matter of more synapses between more neurons.

[9] It is important to note that the same conclusions are forced on us by results elsewhere in neuroscience. Indeed, they reinforce the conclusion about how the brain stores

2. The Darwinian argument for eliminativism

The most powerful philosophical argument for eliminativism that has emerged over the last few decades is due to Darwin, and has been most visibly developed by Jerry Fodor [1990, 2009], though not with an eliminativist agenda.[10]

Physicalist antireductionism needs an account of how a clump of matter, the brain as a whole or more probably a "population" of thousands of neurons wired together into a circuit, has unique propositional content. To do this it needs to show how a clump of matter—a token neural circuit--can be *about* some other thing in the universe.

The best resource, perhaps physicalism's only resource, for explaining how intentionality emerges and what it consists in has to be Darwin's theory of natural selection. There is one huge reason for supposing so. Behavior, including verbal behavior, that is putatively guided by intentional states is purposive, goal directed, it is quintessentially a matter of means aimed at ends. Such purposive behavior inherits its purposiveness from the brain states that drive it. It is why the intentionality of the noises and the marks we make is derived from the original intentionality of neural circuits. But there is only one physically possible process that builds and operates purposive systems in nature: natural selection. That is why natural selection must have built and must continually shape the intentional causes of purposive behavior. Accordingly, we should look to Darwinian processes to provide a causal account of intentional content. That makes teleosemantics an inevitable research program.

Teleosemantics's stock example of how Darwinian processes build intentional content in neural circuitry is the frog's purposive tongue snapping to feed itself flies. The neural circuitry in the frog that produces fly snapping has been tuned up phylogenetically

---

information that emerges from the study of declarative memory. The same detailed account can be given for vision. The visual system conveys information from the retina to the lateral genioculate nuclei, from them via optic radiations to the striate cortex, and from it to afferent behaviors. What neuroscience has discovered is that the visual system is a complex collection of physical-feature "detectors"—sets of cells, neural circuits that produce specific outputs for specific physical inputs, and which combined together produce the beautifully adaptive behavior of a sighted creature behaving with exquisite appropriateness to its environment. It is this appropriateness that impels us to attribute contentful mental states to many creatures and most of all to linguistic ones. But neuroscience has no need of such attributions. There is good reason to conclude that the neural circuits which carry the information we report as propositional are, like the sensory circuits, highly specialized in the features of the world that they store, and that that it is the combination of their effects during retrieval that give the impression that we store whole propositions in memory. Some of the evidence comes from the discovery that information is distributed from the hippocampus into regions specialized to store information from distinct sensory modalities.

As noted below, what is known about conscious awareness also reflects the same character as memory and visual perception.

[10] Fodor's argument was prefigured in Rosenberg, 1986a and 1986b, and employed with the specific aim of advancing eliminativism.

by natural selection and ontogenetically, developmentally, by learning, via the law of effect—operant conditioning Darwinism's chip off the old block.[11] Teleosemantics claims that the neural circuitry's intentional content consists in those phylogenetic and ontogenetic facts about it.

The problem facing teleosemantics is indeterminacy of intentional content. The most exquisite environmental appropriateness of the behavior produced by some neural circuit's firing won't narrow down its content to one unique proposition. This is something that Quine noted under the label of the "indeterminacy of translation". [12] Jerry Fodor labeled this indeterminacy the "disjunction problem" and ever since many writers have used it as a stick with which to beat all causal theories of content.

In the actual environment in which frogs evolved, and in the actual environment in which this frog learned how to make a living, the neural circuitry that was selected for causing the frog's tongue to snap at the fly at x, y, z, t is supposed to have the content "Fly at x,y,z,t." But phylogenetic and ontogenetic Darwinian processes of selection can't discriminate among indefinitely many other alternative neural contents with the same actual effects in tongue snapping behavior. It's now famous that there is no way any teleosemantic theory can tell whether the content of the relevant frog's neural circuit is "Fly or black moving dot at x,y,z,t," or "fly or bee bee at x,y,z,t." or any of a zillion other disjunctive objects of thought, so long as none of these disjuncts has ever actually been presented to the fly. Whence the name, "disjunction problem."

Any naturalistic, purely causal, non-semantic account of content will have to rely on Darwinian natural selection to build neural states cable of having content. This is what teleosemantics seeks to do. But that is exactly what a Darwinian process cannot do.

The whole point of Darwin's theory is that in the creation of adaptations, nature is not active, it's passive. What is really going on is environmental filtration—a purely passive and not very discriminating process that prevents most traits below some minimal local threshold from persisting. Natural selection is *selection against*. As Fodor might put it, Darwin doesn't care which traits get past the filter, including all the bizarre disjunctive traits any student of Nelson Goodman can come up with. Darwin only cares about which traits can't. He and his theory have no time for or need of selection-for. His theory gives pride of place to selection-against. This is not a defect, weakness, oversight or problem of the theory. It is arguably its great strength. Literal *selection for* requires foresight, planning, purpose. Darwin's achievement was to show that the appearance of purpose belies the reality of purposeless, unforesighted, unplanned mindless causation. All adaptation requires is selection against. That was Darwin's point. But the combination of blind variation and selection-against is not possible without disjunctive outcomes. What

---

[11] Dennett, "Why the law of effect won't go away," *Brainstorms*, Cambridge, MIT Press, 1987. For these purposes the frog turns out to be a bad example, since it's close to impervious to operant conditioning. But the example has never been changed to reflect this fact.

[12] It's not as though this problem of indeterminacy escaped the notice of teleosemanticists. Dennett already noticed it in *Content and Consciousness* [1969], though his preferred animal companion was a dog. He detected the indeterminacy problem but he didn't solve it.

Fodor describes as Darwin's disjunction problem is its main achievement![13]

It is important to see that 'selection-against' isn't the contradictory of 'selection for.' Why are they not contradictories? That is, why isn't selection-against trait *T* just selection for trait *not-T*? Simply because there are traits that are neither selected-against nor selected-for. These are the neutral ones that biologists, especially molecular evolutionary biologists, describe as silent, swithced off, junk, non-coding, etc. 'Selection for' and 'selection-against' are contraries, not contradictories.[14]

It is clear that after 50 years or so of trying to come up with a purely causal theory of psychological content that is completely semantics-free, no one has yet succeeded. And that includes Fodor's own beloved asymmetrical causal dependence theory. [15]

---

[13] To see how the process that Darwinian selection-against works in a real case, consider an example: two distinct gene products, one of which is neutral or even harmful to an organism and the other of which is beneficial, which are coded for by genes right next to each other on the chromosomes. This is the phenomenon of genetic linkage. The traits that the genes coded for will be coextensive in a population because the gene-types are coextensive in that population. Mendelian assortment and segregation don't break up these packages of genes with any efficiency. Only crossover, the breaking up and faulty re-annealing of chromosomal strings or similar processes can do this. As Darwin realized, no process producing variants in nature picks up on future usefulness, convenience, need, or adaptational value of anything at all. The only thing mother nature (a.k.a. natural selection-against) can do about the free-riding maladaptive or neutral trait, whose genes are riding along close to the genes for an adaptive trait, is wait around for the genetic material to be broken at just the right place between their respective genes. Once this happens, then Darwinian processes can begin to tell the difference between them. But only when environmental vicissitudes break up the DNA on which the two adjacent genes sit, can selection-against get started—if one of the two proteins is harmful.

Here is Darwinian theory's disjunction problem: the process Darwin discovered can't tell the difference between these two genes or their traits until cross-over breaks the linkage between one gene, that is going to increase its frequency, and the other one, that is going to decrease its frequency. If they are never separated, it will remain blind to their differences forever. What is worse, and more likely, one gene sequence can code for a favorable trait—a protein required for survival, while a part of the same sequence can code for a maladaptive trait, some gene product that reduces fitness. Natural selection will have an even harder time discriminating these two traits.

[14] This feature of natural selection, that it operates on populations to change frequencies by filtering against as opposed to operating on individuals by selecting for adaptations was first made in Sober (1984). The point is quite compatable with his more familiar distinction between 'selection of individuals' and 'selection for properties'. Cf Sober, 1984, 3.2, and 5.2.

[15] Adams, F. and Aizawa, K., "Fodorian Semantics," in S. Stich and T.Warfield (eds.), *Mental Representations*, Oxford: Basil Blackwell, 1994, pp. 223–242, and
Adams, F. and Aizawa, K., "'X' Means X: Fodor/Warfield Semantics," *Minds and Machines*, 4 (1994): 215–231.

Physicalism dictates that psychological states and processes that have intentional content, are just "upgraded neural states" that track the proximate and non-proximate environment with a discriminating enough sensitivity to qualify as representations of particular states of affairs. What counts as 'discriminating enough sensitivity' is relative to the function of the neurological structures that embody the representation. Since (*pace* Fodor 2010) functions are selected effects that already makes teleosemantics the only possible candidate for a theory of content that is itself intentionality-free, that satisfies the physicalist demand that intentional content be upgraded nonintentional content, on pain of begging the question of how intentionality is possible.

Apply these features of the process Darwin discovered to the way neural circuits acquire content: first there is a phylogenetic, evolutionary process that builds neural circuitry and its connections. It selects against circuitry that fails to perform functions required for the organism's survival and reproduction. In circumstances of strong competition, ones in which the bar to survival is set high, this results in neural circuits very finely attuned to their environments. In the case of frogs, neural circuits that send the tongue snapping in even very slightly inaccurate directions are strongly selected against. Whence comes the informational content we ascribe to the circuits which have survived selection against: 'Fly at x,y,z, t.' But of course the process has been unable to discriminate those circuits from ones that cause tongue snapping at disjunctive prey such as 'flie or beebees' or 'flies or black spots on screens in frog's visual field.' We could of course intervene in the course of natural selection to select against neural circuits that have these latter contents, but there are indefinitely many of them and we will never be able to narrow down content to only one disjunct.[16]

Move now from phylogentic to ontogenetic processes. Frogs cannot learn much at all, since they are not subject to substantial operant conditioning, but rats and humans can. Operant conditioning is also a matter of selecting-against. If it were a matter of selecting for, it would lose all its interest as a nonteleolgical account of learning. Operant conditioning over a course of training enables rats to learn certain distinctive behaviors. It does so through a process of feed back in the rat's brain that builds neural circuitry of exactly the same sort as is built by classical conditioning in the sea slug. Teleosemantics bids us attribute propositional content to these circuits, in particular descriptions of the transient envrironment that makes the behavior the neural circuitry produces 'appropriate,' i.e. rewarded. Operant conditioning works by bulding any and every neural circuit that shares a reinforced effect downstream in whatever behavior that is

[16] There is an equally daunting proximal/distal indeterminacy problem that also undermines telesematics' prospects of identifying unique propositional content in neural circuitry. Is it the stimulation in the visual cortex to which the tongue snapping neurons respond, or is it to something further upstream, say the retinal excitations, or is it the photons bouncing off the fly's body, or is it the shape of the fly or its motion, or some combination of them, or the fly itself, or the fly plus the ambient environmental conditions that make it available, or some other factor. As in the disjunction problem, there are indefinitely many links in the causal chain from external sources to the switching on of the right neural circuitry which are equally strongly selected for—i.e. not selected against, as the "referent," "subject," "topic" of the neural circuits' 'content.'

reinforced. Since the behavior doesn't narrow down the upstream causes of the neural circuitry, it cannnot ever narrow down neural content to a unique disjunct.

When it comes to building content teleosemantics is the only game in town since Darwinian natural selection is the only way to get the appearance of purpose wherever in nature it rears its head, and that includes inside the brain. If frogs are hard wired to snap tongues at flies, we have to treat the neural content (fly at x,y,z,t) as a matter of Darwinian shaping of the relevant neural circuits that control frog tongue flicking. In more complex organisms, natural selection first hard wires a capacity to carry information; then learning—classical and operant--shapes the actual informational content of neural circuitry.

If teleosemantics is the only game in town, and if it can't solve the disjunction problem, then the right conclusion is to deny that neural states have as their informational content specific, particular, determinate statements which attribute non-disjunctive properties and relations to non-disjunctive subjects, Thought really is much less determinate than language lets on. The denial that frogs, or for that matter, humans think about *flies*, instead of some (never to be expressed in words) disjunction of *flies or ... or ...* is one that we should take with the utmost seriousness. The disjunction problem is not an objection to teleosemantics. It's a fact of life for biological creatures like us.

3. Consciousness and the introspective illusion of intentionality

50 years of neuroscience have given us ample reason not to trust consciousness or introspection, at least when it comes to developing a theory about the nature of cognition, perception, or emotion for that matter. The way all three of these brain processes manifest themselves in consciousness are symptoms we need to explain, not guideposts on our way towards explanations of how the mind works. Consciousness presumably has a function, in fact, almost certainly more than one. It is too prominent a fact about us not to have emerged and been shaped by natural selection to solve some, probably several "design problems." But exactly what they are and how consciousness disposes of them is not yet known, and will not be revealed by introspection. Meanwhile, the eliminativist cannot take introspection seriously as the basis of a theory that competes with findings and theories in neuroscience.

Yet the chief source of conviction that thought must have intentionality and for that matter unique propositional content, is introspection. It is this unshakeable conviction that is the source of many of the allegations that eliminativism is incoherent. When I look into my self, I know with Cartesian certainty that my thoughts are mainly expressed in sentences and sentence-fragments, silent versions of what I speak, and that these sentences express propositions about the world and myself. When I think that my thoughts are not about anything because there is no "aboutness" or intentionality, I am consciously doing exactly what I claim can't be done: thinking about something. *Reductio ad absurdum*.

No one has advanced the argument for the intentionality of consciousness more explicitly of late that Horgan and Tienson [2010]. What is breathtaking to the eliminativist about this argument that consciousness is sufficient for, and indeed necessary for intentionality, is its question-begging reliance on nothing but introspection. If, as eliminativists hold, the first person point of view is not a reliable source of scientific findings, arguments for intentionality from phenomenological awareness are unavailing.

Add to this what neuroscience can already tell us about neural circuitry and there remains little reason for the neuroscientist to take this argument for the intentionality of thought seriously. This makes arguments for the reality of intentional content from introspection into analyses of the phenomenological origins of an illusion.

> Horgan and Tienson advance the following three theses:
> *The Intentionality of Phenomenology*: Mental states of the sort commonly cited as paradigmatically phenomenal (e.g., sensory-experiential states such as color-experiences, itches, and smells) have intentional content that is inseparable from their phenomenal character.
> *The Phenomenology of Intentionality*: Mental states of the sort commonly cited as paradigmatically intentional (e.g., cognitive states such as beliefs, and conative states such as desires), when conscious, have phenomenal character that is inseparable from their intentional content.
> *Phenomenal Intentionality*: There is a kind of intentionality, pervasive in human mental life that is constitutively determined by phenomenology alone. [Italics in original]

They write, "We argue for the three theses…, in part by way of introspective description of actual human experience. If you pay attention to your own experience, we think you will come to appreciate their truth." They say "in part" but their arguments are solely by way of asking the reader to conduct introspective thought experiments.[17]

What is important for eliminativism is Horgan and Tienson's claim that thinking about things, having thoughts with propositional content, has a qualitative, phenomenal feel to it that makes its aboutness undeniable:

> Consider, for example, an occurrent thought about something that is not perceptually presented, e.g., a thought that rabbits have tails. Quine notwithstanding, it seems plainly false—and false for phenomenological reasons—that there is indeterminacy as to whether one is having a thought that rabbits have tails or whether one is instead having a thought that (say) collections of undetached rabbit parts have tail-subsets. It is false because there is something that it is like to have the occurrent thought that rabbits have tails, and what it is like is different from what it would be like to have the occurrent thought that collections of undetached rabbit parts have tail-subsets.

Horgan and Tienson conclude from this thought experiment that "the phenomenology of these kinds of intentional states involves abstractable aspects which themselves are

---

[17] One set of thought experiments leads to the conclusion that "The full-fledged phenomenal character of sensory experience…involves complex, richly intentional, total phenomenal characters of visual-mode phenomenology, tactile-mode phenomenology, kinesthetic body-control phenomenology, auditory and olfactory phenomenology, and so forth—each of which can be abstracted (more or less) from the total experience to be the focus of attention. This overall phenomenal character is thoroughly and essentially intentional. It is the what-it's-like of being an embodied agent in an ambient environment—in short, the what-it's-like of being in a world." From a purely introspective point of view, this conclusion is hard to argue with. But introspection cuts little ice with eliminativists.

distinctively phenomenological." In fact, according to their introspections, it is the unique propositional content of thought that remains constant over changes in attitude:

> For example, if one contrasts wondering whether rabbits have tails with thinking that rabbits have tails, one realizes that there is something *common* phenomenologically—something that remains the same in consciousness when one passes from, say, believing that rabbits have tails to wondering whether rabbits have tails, or vice versa. It is the distinctive phenomenal character of holding before one's mind the content *rabbits have tails*, apart from the particular attitude type—be it, say, wondering, hoping, or believing. This aspect of the overall phenomenology of intentionality is *the phenomenology of intentional content*."

These are not arguments that will have any force for the eliminativist. In fact, they are powerful and remarkably clear expressions of the illusions that introspection fosters on us and that make eliminativism so difficult to take seriously.

What eliminativism needs is a diagnosis of exactly where this powerful illusion of intentionality in conscious thought comes from. When we begin looking for the sentences in thought that have the original intentionality the first and best candidates are tokens moving across our consciousness when we think. The model of content-conferring acts of conscious thought is that forming the thought that the cat is on the mat is what gives content to the resulting utterance, 'the cat is on the mat'. The tokens of silent speech or mental image sequentially playing across consciousness have content or meaning. The causal pathway to the tongue or hand carries this content to speech or writing. But if these tokenings are just the switching on and off of neural circuits, and neural circuits have no propositional content, then the information conscious thought carries can be no more contentful than the information non-conscious thought carries. Consciousness is just another physical process. If physical processes can't by themselves have or convey propositional content, then consciousness can't either.

To see the problem, lets adopt for the *nonce* a global workspace theory of consciousness [Baars, 1997]. According to this theory consciousness take place in a global work place which a large number of other non-conscious cognitive modules compete temporarily to occupy: aspects of perception, problem solving, planning, language understanding and production. These modules operate in parallel, and whichever gains temporary access to the global workplace broadcasts its information content to the other modules, presumably via its presence in conscious awareness. In effect occupancy of the workplace by one of the modules is what conscious attention consists in.[18] There is increasing neurological evidence (Dehaene & Naccache, 2001; Baars, 2002), including a good deal of neuroimaging data, for the theory that neural circuitry realizes this architecture. Though it would be the product of massive parallel processing, the information flow through the global workspace is serial and the coherent

[18] There is independent evidence for the distributed character of attention: when subjects consciously attend to items in visual or auditory fields, the signature neural correlates occur at the parts of the brain where the earliest, lowest level processing of sensory input arrives. Attention and awareness are distributed processes and not centralized ones, while the molecular biology of both appear to be the same as that of the neural circuitry in the rest of the brain.

outcome of some very complicated set of computational processes. The global workplace model has much to recommend it as the beginnings of a theory of the functional or causal role of consciousness. But for our purposes the sketch suffices to show that just locating the causal role of the neural processes constituting conscious experience cannot help confer content on speech or on brain states for that matter.

Like the rest of the brain, the global work place is a network of neural circuits, operating on exactly the same principles as all the others. Suppose the global workplace's role in bringing about speech is to be the scene of a serial sequence of tokens, markers, silent phonemes or word sounds, perhaps visual shapes. The question immediately arises as to what gives these tokens the content or meaning that they eventually accord to spoken or written tokens. Content cannot be conferred upon conscious tokens in virtue of their composition out of neural circuitry or its firing, as we have already excluded the possibly that neural circuits carry information symbolically, let alone sentential. The silent sounds and images in consciousness are themselves fully physical. Whatever it may be like to think 'the cat is on the mat' these qualitative aspects of conscious thought can't convey intrinsic intentionality to the thought itself, if they are material aspects of neural circuitry. And they can't do it if they are nonmaterial either, unless dualism is right and comes equipped with an adequate theory of non-physical causation.

The silent "sound" tokens and images in our consciousness are in exactly the same boat as the spoken tokens and inscriptions in public speech. They are the result of Darwinian selection on neural circuitry that makes possible coordination, collaboration, and cooperation among big-brained primates. In its broad outlines the natural history of language is well understood. What humans especially needed, once they found themselves on the bottom of the African savanna food chain, was a means to defend themselves against mega fauna, then to scare them off their prey so humans could scavenge it, and finally to attack the mega fauna themselves. The co-adaptational cycle of improving coordination and increasing protein nutrition produced signaling, incipient pidgins, creoles, and eventually full-blown public language, along with an unavoidable accompaniment in conscious thought. But neither language nor consciousness required nor came equipped with unique propositional content or individual meanings for the mental tokens--terms and predicate--that are supposed to be combined to express them.

Eliminativists treat intentionality—original and derived-- as a myth that emerges from the earliest attempts to explain how spoken and written signs become symbols. Its mythic status is clear once we see that there are no symbols, just signs.

To this analysis eliminativists may adapt an argument of Horgan and Tienson, one which shows that intentionality of thought is a figment of its linguistic character. They write:

> [T] he what-it's-likeness of intentionality that we are talking about…. attaches to awareness of …words qua contentful; *it is the what-it's-like of hearing or saying those words when they mean just that: that rabbits have tails*. So the basic point holds:… if thinking …involve[s] auditory imagery, the auditory imagery would be intentionally loaded in the experience, not intentionally empty. [Italics added]

Horgan and Tienson invoke a particularly attractive phenomenological thought-experiment from Galen Strawson [1994] that is supposed to show the intentionality of mental tokens. But it is sufficiently rich in detail that it enables the eliminativist to identify clearly the source of the illusion of intentionality in conscious thought.

> As they write, Strawson invites us to
> consider the phenomenological difference between hearing speech in a language that one does not understand and hearing speech in a language that one does understand…. At a certain relatively raw sensory level, their auditory experience is phenomenologically the same; the sounds are the same, and in some cases may be experienced in much the same way *qua sounds*. Yet it is obvious introspectively that there is something phenomenologically very different about what it is like for each of them: one person is having understanding experience with the distinctive phenomenology of understanding the sentence to mean just what it does, and the other is not.

The tendentious character of this description is easy to recognize. A more neutral description makes the eliminativist's point clearly; the phenomenal difference here is a matter of differences in the sequence of silent 'sound' tokens that flit across consciousness, along with sensations and feelings that pass through it. In the case of a listener who speaks the same language, the tokens, images and other mental items are ones associated with memories and environmentally appropriate verbal behavior. In the case of a listener who does not speak the language, the items usually include ones associated initially with recall of sounds heard in the past, verbally expressed mental queries, and then with a feeling of annoyance owing to the incoherence of the spoken noises with the hearer's thoughts. The difference is just a difference in the order and connection of ideas, unless of course we are prepared to accept blatantly question-begging descriptions of the difference. "Consciously understanding meanings" is not some special intentionally freighted achievement. It's having a sequence of tokens in consciousness that bring about a sequence of environmentally appropriate verbal behaviors.[19] The sequence of tokens in consciousness and its behavioral accompaniments is simply different from those of a person who does not speak the relevant language. It's the train of images and tokens in consciousness that tricks us into the whole common sense theory of intentionality and aboutness.

Consciousness is no more capable of grounding the attribution of unique propositional content to neural circuitry than is the behavior that it accompanies and perhaps even causes.

---

[19] Horgan and Tienson give the eliminativist a nice example to illustrate the eliminativist's treatment of conscious understanding that the reader can use:

> Consider, as a similar example for a single speaker, first hearing "Dogs dogs dog dog dogs," without realizing that it is an English sentence, and then hearing it as the sentence of English that it is. The phenomenal difference between the experiences is palpable. (If you do not grasp the sentencehood of the "dogs" sentence, recall that 'dog' is a verb in English, and compare, "Cats dogs chase catch mice.")

The eliminativist observes that the first time you read the five almost identical inscriptions, 'dog' and 'dogs' no set of experiences, images, pictures other than the inscriptions played across your consciousness. The second time a quite different set did so, with different effects. The differences in mental items are all there is to the illusion of propositional content.

4. Dealing with eliminativism's incoherence problem while explaining away the illusion of intentionality.

There is a great deal of science that stands behind eliminativism and underwrites its claim that neural circuitry does not encode sentences (or anything like them) expressing or representing unique propositions. Neuroscience will eventually get around to providing a correct account of how the brain acquires, stores, and employs information. When it does so this account will be written down in sentences that seem to express true propositions about how the brain does it. That is the real problem for eliminativism. For eliminativism bids us recognize that these sentences will have no meaning, express no true propositions, and so tell us noting about how the brain works. If eliminativism is right, it can't be expressed, expounded, defended, or adopted. In fact the same goes for all the science that stands behind it. This is the real *Reductio Ad Absurdum* of eliminativism.[20]

How much of the force of this objection can the eliminativist reduce while consistently maintaining that neither expressed sentence tokens nor brain states that bring them about have propositional content? One approach that can be ruled out is some sort of instrumentalism about propositional content. It is no solution to eliminativism's problem to adopt an intentional stance, one that instrumentally *interprets* neural circuitry or its effects in speech and writing as contentful. The eliminativist cannot help herself to an interpreter to take up the intentional stance, to merely use, without endorsing, the hypothesis that other people and animals have brain states with propositional content. That way lies regress. For eliminativism is the thesis that literal interpretation never happens: interpreting something is translating it, putting it in other words, bringing it under a description, treating it as a hypothesis. To do any of these things our neural circuitry would have to contain sentences that express the interpretations.

Eliminativism's problem is not that it denies brains store information, nor even the problem of explaining at least schematically how they do so. Neuroscience has begun to give it the detailed answers to the questions about the various ways in which neural circuitry is organized to acquire, store and deploy information, including the information it needs to enable the body to produce language. None of these ways the neural circuitry does its job requires it to store unique propositions or disjunctions of them.

Recall the point made at the outset. Eliminativism does not deny that one main way in which information is conveyed between brains is via spoken and written language. Speech and writing do this. They carry information from brain to brain, and they appear to have content, to express unique propositions. But that can't be the way they carry information, because if they did, then the neural circuits, from which and to which the information is communicated, would also have to carry this information in the same way. So, how can sentences carry information without expressing propositions?

To begin to explore a solution to the eliminativist's problem consider a simple alternative, one not unfamiliar in philosophy's recent past, which employs the metaphor of a map. Start with a political map of Europe. It is easy to "read off" from this map an

---

[20] There are a variety of alternatives on this *reductio,* or self-refutation, pragmatic contradiction objection, that have been advanced against eliminativism. The best of these still seems to be Lynn Rudder Baker's *Saving Belief* (1987). The version articulated above seems the most serious version of this objection.

indefinitely large number of distinct pieces of information that the map stores nonsententially: "Paris is east of London", "London is west of Paris," "Paris is a city," "Paris is a national capital," and an indefinitely large, perhaps infinite number of other such true sentences. The set of such sentences could also be used by someone who hears or sees them to draw a political map of Europe, one which, given enough time, asymptotically approaches the first map in its informational content. Yet, neither of the two maps is a set of sentences expressing unique propositions. How much of this metaphor can be converted to a literal claim about how the brain stores information and how language communicates it without actually having propositional content? Quite a lot.

Of course eliminativism cannot help itself to the literal conception of a map. Maps are representations just like sentences. One way to see this is simply to think about the various projections in which maps can be drawn, the different kinds of maps—political, topographic, demographic. Each requires a "key", a set of instructions about how to *interpre*t the map. If the arrangement of neural circuits in the brain maps the world, reality, the brain's various environments, it can't do so by bearing a relation to them that, like a map's relation to what it maps, is mediated by *interpretation* of the map. That way lies regress or circularity—what interprets the neural circuits that interpret the map? The relationship between neural circuits and the world that they "map" must be some sort of physical relation, more likely, not just one physical relation but many different ones, which vary depending on the features of the world various neural circuits "map."

Discovering how the behavior of neural circuits "maps" their causes and their effects is at the top of Cognitive neurosciences' agenda. By uncovering the details neuroscientists have begun to solve eliminativism's incoherence problem. This work was not undertaken with a view to solving the eliminativist's coherence problem. It was undertaken in order to figure out exactly how neural circuits store information. But attracts the attention of any philosopher dissatisfied with teleosemantics inability to provide a thoroughly non-intentional grounding of cognitive content. To that extent it will be unsurprising that the eliminativist may be able to make use of this research to deal with the incoherence problem.

These neuroscientific discoveries about neural information-storage have encouraged the development of what philosophers call "structural resemblance theories," which invokes a relationship between the physical structure of a neural circuit and the object or state of affairs it is said to carry information about. Accurate maps bear structural relations to the geography they map: the spatial relations among the marks on the map preserves the spatial relations among the geographic items they are interpreted by us as mapping, and they do so independent of any interpretation we provide. The structural relation here is "first order:" spatial relations on the map that are structurally similar to spatial relations in the world. But there are also "second order structural similarities:" These are the relationships many measuring instruments, especially dials on dashboards, exploit. The simplest and perhaps oldest is the second order relationship exploited by a pan balance scale: the relationship between the mass of an object and spatial displacement of the pointer on the scale is structural, but it is not a matter of spatial relations reflecting spatial relations; rather spatial relations—the pointers movements, bear a structural relation to differences in mass.

Structural resemblance is then a relation that obtains between two things when their respective parts stand to one another in one or more physical relations. "Second order

resemblance" obtains when the two sets of components of each thing share the same abstract relationship to one another. As illustrated below, in the case of neural circuits that contain information or misinformation about the world, sharing one or a small number of these perhaps multitudinous physical relations will be the most relevant to what information or misinformation it contains. And of course natural selection has disposed organisms, in this case humans, to behave in ways very finely adapted to exploiting the particular structural identity between components of the neural circuitry and what it bears an informational relation to.

Experiments in neuroscience uncover such structural relations between environmental processes and the informational content of the neural circuits they cause and between the structural character of the neural circuits and the environmentally appropriate behavior they bring about. In the 1980's, experiments with macaque monkey, have isolated the structural resemblance between input vibrations the finger feels, measured in cycles per second, and representations of them in neural circuits, measured in action-potential spikes per second [Mountcastle, Steinmetz, and Romo, 1990]. This resemblance between two easily measured variables makes it unsurprising that they would be among the first such structural resemblances to be discovered. Macaques and humans have the same peripheral nervous system sensitivities and can make the same tactile discriminations. Experiments on macaques have shown how the structural representations of different input stimuli are computationally compared in the macaque brain, and show how they cause output behavior that reflects the comparisons when the macaque is subject to operant reinforcement for the discriminations. These were baby steps in deciphering the purely physical discriminations that perception, memory, cognitive processing and motor control consist in. They and subsequent research into the neural processing involved in much more complex have increasingly vindicated a structural resemblance approach to how information enters the brain, is stored, and deployed.

It is obvious how structural resemblance theory lends itself to theories of information as causal covariance [*sensu* Dretske, 1981], and theories that accord neutral circuitry the function of storing such information. The causally covariance of neural circuitry with any of its prior causes and future effects, will include inputs that make its outputs environmentally appropriate and so accord the neural circuitry adaptational functions. Note, it can do all this without these neural circuits having propositional content. That is one reason philosophers hoping for, or challenging the possibility of, a physical account of semantic information storage in the brain have been dissatisfied with such theories of information as causal covariance.

Eliminativists will see the causal covariance theory of information as a cup more than half full. That is, not holding with intentional content to begin with, they will not require that an adequate theory of how the brain stores information accord it storage of propositional content. Thus, linguistic expressions, sounds or inscriptions that move information from brain to brain by the use of sounds or marks that must be taken up in a temporally extended process such as reading or hearing, convey information by rearranging neural circuits in one head to bear new relations of (mainly second order) structural similarity to ones in another head.

It is not hard to see how a structural resemblance theory can come to the aid of the eliminativist in the project of blunting the charge of incoherence. Start with the simple

case of how information about the frequency of a tactile stimulation in the finger gets stored in the neural circuitry and eventually results in a sentential vocalization by a human subject: "the frequency of stimulation has increased." The eliminativist and the neuroscientist take this vocalization and its apparent propositional content seriously as a reliable effect of the information stored, without however treating its apparent semantic or syntactic structure as indicative of the way in which the information is stored in the relevant neural circuitry. The semantic and syntactic features of the vocalization are apparent, and not real, of course since to be real they require prior original intentionality in their causes—the neural circuitry of the brain.

It will be an important project for cognitive neuroscience, and especially neuro-psycholinguistics, to develop a theory of why and how information stored non-sententially in the brain is communicated between brains by processes—speech and writing—that have a temporal and spatial structure, one that gives rise to the illusion of syntax and semantics. This theory would presumably also be relevant to understanding how nonsententially stored information gives rise to silently sounded sentences in conscious thought. One of the special constrains under which the development of such a theory will operate is the reflexive fact about the theory that it will have to be couched in terms of the very illusion it explains: sentences in a spoken or written language.

One of the adequacy conditions on a neural account of speech production will be that it provide a rough translation manual, enabling us to infer from sentences that speakers/writers "sincerely"[21] produce back to the neural circuitry that (nonsententially) stores the information causing the vocalization or inscription. Presumably the theory would only be pressed into service in psychophysics laboratories and perhaps neurological diagnosis and treatment. What we already know about the widely distributed character of information storage in the brain and what we surmise about the multiple realizability of cognitive states by neural circuitry provide grounds to expect only a very inexact translation manual. If neural circuits carry information in anything like the way maps do, then the translation manual will face even greater difficulties, owing to the potentially infinite number of sentences required to convey all the information carried by any map.

So, if eliminativism is correct the translation manual will in most cases remain very rough and approximate in the guidance it offers to exactly what information is carried in neural circuitry. In fact, exact and accurate translation would be excellent evidence that neural circuits do carry information in sentences, or some "data structures" that can be systematically mapped on to them (which would amount to sentential information storage).

We can apply this machinery to make "sense" of eliminativism in terms of the sentences the eliminativist speaks or writes. When we say that eliminativism is true, that the brain does not store information in the form of unique sentences, statements, expressing propositions or anything like them, there is a set of neural circuits that have no trouble coherently carrying this information. There is an as yet unavailable but in principle possible translation manual that will guide us back from the vocalization or inscription eliminativists express to these neural circuits. These neural structures will differ from the neural circuits of those who explicitly reject eliminativism in ways that

---

[21] Scare quotes owing to the intentionality of sincerity.

presumably our translation manual may be able to shed some light on: giving us a neurological handle on disagreement and on the structural differences in neural circuitry, if any, between asserting p and asserting not-p when p expresses the eliminativist thesis expressed above.

Is this enough to solve the eliminativist's apparent self-refutation problem? Is talk about translation manuals anything more than an eliminativist circumlocution for all the things non-eliminativists can already say using the vocabulary of propositions, truth, reference, satisfaction, and all the rest of the machinery of intentional content? Circumlocution is all we need to avoid the charge of incoherence.

In adopting this set of alternative descriptions of what is going on in thought, eliminativists are embracing a venerable strategy in philosophy. It is one first explicitly advanced by Bishop Berkeley, when he invited us to "Think with the learned, speak with the vulgar." My aim here has been to make eliminativism at least coherent, given that it is evidentially compelling. Doing that requires that we explain how we can store the information speech and writing express variously as the denial that there are beliefs, meaningful expressions of beliefs or even theses to be believed. Thinking with the learned, especially the neuroscientist, we recognize that information the brain acquires, stores and employs, doesn't come in sentences or anything like them. But speaking with the vulgar, we can accept that brains convey this information in sounds and inscriptions that confer the illusion that there are propositions that give the content of the sentences speech and writing convey. And that goes for all the sentences in this paper.

Alex Rosenberg
Duke University

References

Baars, 1997 *In the Theater of Consciousness: The Workspace of the Mind*. New York: Oxford University Press.

Baker, L. 1987. *Saving Belief*. Princeton, Princeton University Press.

Dretske, F., 1981, *Knowledge and the Flow of Information*, MIT Press.

Dretske, F., 1988, *Explaining Behavior*, MIT Press

Fodor, J. 1990, *A Theory of Content and Other Essays*, MIT Press

Fodor, Jerry, *What Darwin Got Wrong*, with Massimo Piattelli-Palmarini, Farrar, Straus and Giroux, 2010

Horgan and Tienson 2010,
[http://www.u.arizona.edu/~thorgan/papers/mind/IPandPI.htm#_edn1].

Kandel E, Bailey, C, Bartsch, D 1996, "Toward a molecular definition of long-term memory storage", *Proceedings of the National Academy of Science* 93 (24): 13445–13452,

Mountcastle VB, Steinmetz MA, Romo R, 1990, "Frequency discrimination in the sense of flutter: psychophysical measurements correlated with postcentral events in behaving monkeys," *Journal of Neuroscience*,10(9):3032-44.

Rosenberg, A, 1986 "Intentional Psychology and Evolutionary Biology Part I: The Uneasy Analogy," *Behaviorism,* 14: 1:5-27

Searle, John 1980, "Minds, Brains and Programs," *The Behavioral and Brain Sciences*.3, pp. 417–424.

Searle, John 1983, *Intentionality: An Essay in the Philosophy of Mind*, Oxford, Oxford University Press

Sober, Elliot 1984, *The Nature of Selection*, Cambridge, MIT Press.

Strawson, Galen 1994, *Mental Reality*, Cambridge: MIT Press.